

①9 BUNDESREPUBLIK
DEUTSCHLAND



DEUTSCHES
PATENTAMT

⑫ **Offenlegungsschrift**
⑩ **DE 44 42 958 A 1**

⑤1 Int. Cl.⁶:
G 06 F 17/14

②1 Aktenzeichen: P 44 42 958.4
②2 Anmeldetag: 2. 12. 94
④3 Offenlegungstag: 5. 6. 96

DE 44 42 958 A 1

⑦1 Anmelder:
SICAN Gesellschaft für Silizium-Anwendungen und
CAD/CAT Niedersachsen mbH, 30419 Hannover, DE

⑦2 Erfinder:
Fiedrich, Sven, 10825 Berlin, DE; Grüger, Klaus, Dr.,
30459 Hannover, DE; Müller, Gerriet, 10997 Berlin,
DE; Schneider, Ulrich, 13593 Berlin, DE

⑤6 Für die Beurteilung der Patentfähigkeit
in Betracht zu ziehende Druckschriften:

DE	39 00 349 C2
DE	34 16 536 A1
US	52 93 330
US	48 21 224

⑤4 Verfahren und Schaltungsanordnung zur Durchführung mehrstufiger Butterfly-Operationen

DE 44 42 958 A 1

Beschreibung

Die Erfindung betrifft ein Verfahren und eine Schaltungsanordnung zur Durchführung mehrstufiger Butterfly-Operationen, wie sie z. B. bei der Transformationscodierung und -decodierung, insbesondere für die Diskrete-Cosinus-Transformation (DCT) verwendet werden.

Technisches Gebiet

Verschiedene Blocktransformationsalgorithmen lassen sich besonders effektiv berechnen, wenn die Berechnung mehrstufig ausgeführt wird. Der Signallaßgraph wird dafür in kleinere, möglichst ähnliche Teilgraphen mit je zwei Ein- und Ausgängen zerlegt, die meist die Form eines Schmetterlings haben und daher Butterfly genannt werden. Üblich sind dabei Butterflies mit jeweils zwei Ein- und Ausgängen. Als interne Rechenoperationen für die Butterflies werden üblicherweise Additionen, Subtraktionen und die vergleichsweise aufwendigen Multiplikationen benötigt. Im Signallaßgraphen werden dabei in jeder Stufe mehrere Butterflies parallel benötigt. Beispiele dafür sind die Fast-Fourier-Transformation (FFT), die ausführlich in "Digital Signal Processing" von A. V. Oppenheim und R. W. Schäfer, Engelwood Cliffs: Prentice Hall, 1975, beschrieben ist, und die Diskrete-Cosinus-Transformation (DCT) und die inverse DCT (IDCT). Weitere Ausführungen finden sich auch in:

[N. Ah med, T. Natarajan & K. R. Rao: "Discrete Cosine Transform", IEEE transactions on Computer, Vol. C-23, S. 90—93, Jan. 1974],

[B. G. Lee: "A New Algorithm to Compute the Discrete Cosine Transform", IEEE transactions on Acoustics, Speed and Signal Processing, Vol. 32, No. 6, S. 1243—1245, Dec. 1984],

[A. Artieri, F. Jutand: "Procédé de détermination de transformée en cosinus discrète", Brevet No. 89 02347, Feb. 1989], [C. Loeffler, A. Lightenberg, G.S. Moschytz: "Practical Fast 1-D DCT Algorithmus with 11 Multiplikationen", IEEE, S. 988—991, 1989],

[M. Vetterli, H. Nussbaumer: "Simple FFT and DCT Algorithms with Reduced Number of Operations", Signal Processing, Vol. 6, S. 267—278, Aug. 1984].

DCT und IDCT werden häufig für die Datenkompression von Videosignalen eingesetzt. Die Verfahren sind in den Standards JPEG, MPEG, "Information Technology — Coding of Moving Pictures and Associated Audio for Digital Storage Media up to about 1.5 Mbit/s", Draft International Standard ISO/IEC DIS 11 172, 1992, definiert. Dabei wird eine zweidimensionale Form der DCT eingesetzt, die in zwei eindimensionale DCT-Transformationen mit einer Länge von 8 separierbar ist.

Das Problem ist auch in "Video Codec for Audiovisual Services at $p \times 64$ kbit/s, Recommendation H. 261", The International Telegraph and Telephone Consultative Committee (CCITT), 1990 näher definiert.

Zur Implementation von Signallaßgraphen für Blocktransformationsalgorithmen stehen im Prinzip mehrere Möglichkeiten zur Verfügung. Der gesamte Signallaßgraph kann direkt in Hardware überführt werden, so daß jede Multiplikation einen festverdrahteten Multiplizierer, jede Addition einen Addierer und jede Subtraktion einen Subtrahierer erfordert. Diese Vorgehensweise führt zwar zu extrem hohen Durchsatzraten, ist jedoch mit vergleichsweise hohem Aufwand verbunden. Das andere Extrem liegt darin, alle Operationen

nacheinander in einer programmierbaren Einheit im einfachen Zeitmultiplex durchzuführen. Ein Beispiel ist auf Seite 265 in dem Datenhandbuch von LSI-Logic ausgeführt (Implementing Fast Fourier Transform Systems with the L64280/81 Chipset. Digital Signal Processing (DSP) Databook. Milpitas: LSI Logic, June 1990, S. 258—268). Dieses Vorgehen ist zwar sehr flexibel und liefert hohe Auslastungsraten, führt jedoch zu niedrigen Durchsatzraten.

Häufig wird hierbei wird daher die Möglichkeit genutzt, die Anzahl der realisierten Butterflies über einen Zeit-Multiplexbetrieb der gewünschten Durchsatzrate anzupassen. Dabei ist in der Regel jeder Verarbeitungstufe eine Butterfly zugeordnet (LSI-Logic, Databook, Seite 261, Fig. 1 und 2). Die Zuordnung kann beispielsweise über einen Projektionsalgorithmus erfolgen. Für den Betrieb werden zusätzlich zu den eigentlichen Butterflies Permutationselemente benötigt, um die Daten in die jeweils vorgegebene, in jeder Stufe unterschiedliche zeitliche Reihenfolge zu bringen.

In den meisten Fällen wird jedoch bei der Konzeption von Verarbeitungskomponenten ein einfaches Systemkonzept angestrebt, indem im Mittel mit jedem Taktschritt ein Datum verarbeitet wird. Jeder Taktschritt kann dabei ggf. auch aus mehreren Teiltakten bestehen. Auch synchron getaktete Zwischenspeicher, die eventuell zwischen mehreren Verarbeitungseinheiten zur Anpassung von Datenformaten benötigt werden, liefern pro Taktschritt meist maximal ein Datum, sofern sie nicht aufwendig unterteilt sind. Deshalb werden manchmal auch Butterflies im Zeitmultiplex für mehrere Stufen gemeinsam eingesetzt. Um eine optimale Auslastung der Verarbeitungseinheiten (Prozessoren) und damit einen möglichst niedrigen Aufwand zu erhalten, muß die Anzahl der Prozessoren an die Durchsatzrate angepaßt werden.

Die aufwendige Implementierung von Multiplizierern kann dadurch minimiert werden, daß die Transformation mit einer möglichst großen Anzahl von Multiplikationen mit festem Koeffizienten ausgeführt wird. Zum Beispiel bei der Kombination von festverdrahteten Addier-, Subtrahier- und Schiebe-Operationen liegt der Aufwand für eine Multiplikation mit festem Koeffizienten erheblich niedriger, als für eine Multiplikation mit variablem Koeffizienten [K. Hwang: Computer Arithmetic, Principles, Architectures and Design. John Wiley & Sons, 1979] und (P. Pirsch: VLSI Implementation Strategies. in: P. Pirsch (Hrsgb.): VLSI Implementations for Image Communications, Elsevier 1993, S. 67].

Die Butterflies arbeiten mit 2 Werten pro Taktschritt meist schneller als die Speicher und die Datenzu- und -abfuhr, die meist 1 Wert pro Taktschritt benötigen. Entsprechende Verhältnisse ergeben sich auch bei der Verarbeitung komplexer Zahlen wie sie beispielsweise bei der FFT benötigt werden. Häufig werden alle n Ebenen zeitlich nacheinander abgearbeitet. Bei optimaler Datenzuführung über angepaßte Zwischenspeicher sinkt dann aber die bei n Stufen erreichte Durchsatzrate auf das $2/n$ -fache des Wertes ab. Für die Erzielung einer optimalen Durchsatzrate mit minimalem Aufwand ist es daher notwendig, die Anzahl der realisierten Butterflies so zu wählen, daß der Wert $n/2$ der Anzahl der zu verarbeitenden Werte je Taktschritt entspricht. Bei ungerader Anzahl n ergibt sich jedoch das Problem, daß $n/2$ nicht ganzzahlig ist. Entsprechendes gilt für vergleichbare Aufgaben, bei denen die optimale Anzahl der Butterflies nicht ganzzahlig ist. In all diesen Fällen ist bislang keine optimale Implementierung möglich.

Dies trifft beispielsweise auf dreistufig implementierte Signallaßgraphen für die 8-Punkte-DCT zu. Wird dabei beispielsweise mit nur einem Butterfly-Prozessor gearbeitet, so wird zwar der Verarbeitungseinheit zeitlich vollständig genutzt, die Durchsatzrate von Speicher und Gesamtsystem ist jedoch auf 2/3 des Maximalwertes gesunken. Werden dagegen zwei, bzw. drei Butterfly-Prozessoren eingesetzt, um eine Durchsatzrate von einem Wert/Takt zu erreichen, so sind dagegen die Prozessoren nicht vollständig ausgelastet.

Aufgabe der Erfindung

Aufgabe der Erfindung war es, ein Verfahren und eine Schaltungsanordnung zur Durchführung mehrstufiger Butterfly-Operationen anzugeben. Der Implementierungsbedarf, insbesondere die Anzahl an Multiplizierern sollte hierbei minimal sein.

Erfindung

Die Aufgabe wird durch das erfindungsgemäße Verfahren nach Anspruch 1 gelöst. Mehrere Stufen einer Transformation werden im einfachen Zeitmultiplex in einer gemeinsamen Schaltung zur Durchführung von Butterfly-Operationen ausgeführt. In der erfindungsgemäßen Schaltungsanordnung nach Anspruch 3 wird ein Rechenwerk zur Durchführung von Butterfly-Operationen für die Stufen gemeinsam genutzt. Die Daten werden nach der Berechnung einer Stufe permutiert und über eine Verzögerungsleitung in das Rechenwerk zurückgekoppelt. Im Zeitmultiplex können somit Berechnung für die Stufen abwechselnd im gleichen Rechenwerk ausgeführt werden.

Die Butterflies in den Signallaßgraphen für die Blocktransmutationsalgorithmen müssen hierbei so geschickt gewählt sein, daß nur in maximal $n-1$ Stufen die vergleichsweise aufwendigen programmierbaren Multiplikationen implementiert werden müssen.

Zur Durchführung von Butterfly-Operationen, z. B. für FFT oder DCT, wird an den Ausgang der Schaltungsanordnung zur Berechnung mehrerer Stufen im Zeitmultiplex ein Rechenwerk zur Durchführung von Butterfly-Operationen und Permutationen geschaltet, das ohne stufenweisen Zeitmultiplex arbeitet. Für die Berechnung der inversen Transformationen wird ein ohne stufenweisen Zeitmultiplex arbeitendes Rechenwerk zur Durchführung von Butterfly-Operation und Permutationen der durch Zeitmultiplex gekennzeichneten Schaltung vorgeschaltet.

Bei einer ungeraden Anzahl n läßt sich so mindestens eine der Butterfly-Verarbeitungseinheiten mit festem bzw. umschaltbaren Multiplikationskoeffizienten realisieren. Diese Verarbeitungseinheit ist zumeist am Ein- bzw. Ausgang der Schaltung angeordnet. Es ist daher möglich, daß $n-1$ Butterfly-Stufen im zweifachen Zeitmultiplex betrieben werden und die verbleibende Stufe separat ohne Zeitmultiplex implementiert wird. Dadurch wird bei den aufwendigen Stufen ein Ausnutzungsgrad von 1 erreicht, während lediglich bei der ersten, bzw. letzten Stufe der Ausnutzungsgrad bei 0,5 liegt. Da die separate Stufe im allgemeinen nur einen Bruchteil des Gesamtaufwands ausmacht, wird so Gesamtausnutzungsgrad von fast 1 erreicht. Entsprechendes läßt sich häufig auch bei anderen Durchsatzraten erreichen.

Die Erfindung wird anhand der Figuren näher erläutert. Es zeigen:

Fig. 1 Beispiel eines geeigneten DCT-Datenflußdiagramms mit drei Stufen.

Fig. 2 Gemischter Multiplex-Normal-Betrieb mit optimaler Ausnutzung der Butterflies.

Fig. 3 Gemischter Multiplex-Normal-Betrieb für IDCT (mit IDCT-Butterflies).

Fig. 4 Permutationsschaltung.

Anwendungsbeispiele

Die Erfindung findet z. B. Anwendung bei den vergleichsweise regulär aufgebauten Transformationen FFT und IFFT, beschrieben in "Digital Signal Processing" (Seite 293, Fig. 6.6, bzw. Seite 304, Fig. 6.17), sowie für einige relativ irregulär aufgebauter DCT- und der IDCT-Implementierungen, wie etwa die auf dem Algorithmus von Lee basierenden Signallaßgraphen für die 8-Punkte-DCT (N. Demassieux, F. Jutand: "Orthogonal Transforms", in P. Pirsch, "Implementations for Image Communications", Elsevier, 1993, S. 217—250, insb. S. 223, Fig. 3).

Bei der FFT wird mit komplexen Zahlen gerechnet. Die benötigten Operationen lassen sich aber auch auf reelle Additionen, Subtraktionen und Multiplikationen zurückführen. Bei den Formen der FFT wird neben Addition und Subtraktion in der ersten bzw. letzten Stufe lediglich die Multiplikation mit dem Wert 1 benötigt, in der zweiten, bzw. vorletzten Stufe die Multiplikation mit der imaginären Zahl i . Beide lassen sich leicht ohne eine Hardware-Multiplizierer realisieren. In der dritten, bzw. drittletzten Stufe wird zusätzlich die Multiplikation mit der Quadratwurzel von 2 benötigt, die dann mit festverdrahteten Multiplizierern realisieren läßt. In diesem Falle kann sich das Herausnehmen von bis zu drei Stufen lohnen.

Auch bei der für Bildcodierungsverfahren häufig benötigten DCT der Länge 8 lassen sich Butterflies so finden, daß nur in den ersten beiden Butterfly-Stufen B1 und B2 die sehr aufwendigen, programmierbaren Multiplizierer benötigt werden (Fig. 1). In der dritten Butterfly-Stufe B3 wird statt dessen lediglich ein sehr einfacher, festverdrahteter Multiplizierer benötigt, der mit der Quadratwurzel von 2 (Faktor $2 \cdot P4$) oder mit der Hälfte der Quadratwurzel von 2 (Faktor $P4$) multiplizieren muß. Die Umschaltung zwischen diesen beiden Multiplikatoren ist daher sehr einfach möglich. Beispielsweise erfordert sie bei Verwendung des üblichen Dualzahlensystems lediglich ein bitweises Verschieben der Multiplikanden. Zusätzlich ist ein einfaches Korrektur-Netz K und eine Skalierung Sk notwendig. Das Korrektur-Netz enthält nur einfache Additionen und kann daher vergleichsweise einfach aufgebaut werden. Die Skalierung für den Faktor $P4/2$ entspricht einem Viertel der Quadratwurzel von 2. Wird eine zweidimensionale DCT (2D-DCT) wie für Bildcodierungsaufgaben berechnet, so wird diese Skalierung zweimal benötigt. Aufgrund der prinzipiellen Linearität der Anordnung lassen sich daher die beiden Skalierungsfaktoren zum Faktor $1/8$ zusammenfassen. Dieser Faktor entspricht einer Zweierpotenz und kann daher in Verbindung mit den üblicherweise eingesetzten Zahlensystemen mit einer festverdrahteten Schiebe-Operation ohne zusätzliche Hardware implementiert werden.

Insgesamt wird also neben dem Aufwand für die Im-

plementierung der Datenzu- und -abführung nur Hardware für die Implementierung der 1., 2. und 3. Butterfly-Stufen mit Permutation und dem Korrektur-Netz benötigt.

Es ist daher möglich, die Verarbeitung erfindungsgemäß so aufzuteilen, daß die beiden ersten aufwendigen Butterfly-Stufen B1/B2 im Zeitmultiplex arbeiten, die letzte, einfach aufgebaute Butterfly-Stufe B3 jedoch direkt mit dem Verarbeitungstakt betrieben wird (Fig. 2). Die Stufe B3 muß dabei auch die für -das notwendige Korrektur-Netz beinhalten. Für die Einführung der Zeitmultiplex-Verarbeitung sind lediglich zusätzliche synchrone Verzögerungen D (Pipeline-Stufen) und Multiplexer M erforderlich, um die Serien/Parallel-Wandlung und die Parallel/Serien-Wandlung durchzuführen. Wird die Stufe 3 im Zeitmultiplex betrieben, so wird die Parallel/serien-Wandlung erst nach dieser Stufe durchgeführt (Fig. 3).

Bei einem derartigen Multiplex-Betrieb wird zumindest der aufwendig aufgebaute Teil der Verarbeitungseinheit vollständig genutzt, ohne daß sich die maximal mögliche Durchsatzrate reduziert wird. Diese Eigenschaft ermöglicht effiziente Hardware-Realisierungen.

Mit angepaßten Butterflies und Umkehrung der Verarbeitungsmodi läßt sich dieses Grundprinzip auch auf die IDCT-Verarbeitung anwenden (Fig. 4). In diesem Fall braucht der für die Butterfly B1 eingesetzte Multiplizierer nur mit festverdrahteten Koeffizienten ohne Zeitmultiplex betrieben werden, während die aufwendigen Einheiten für die Butterflies B2/B3 im Zeitmultiplex betrieben werden. Auch hierbei sind zusätzlich wieder lediglich wenige Verzögerungen D und Multiplexer für den Zeitmultiplex-Betrieb erforderlich.

Zwischen den Stufen müssen die Daten ggf. permutiert werden, damit die Ausgangsdaten für die nächste Stufe in der richtigen Reihenfolge am Eingang der Butterfly-Stufe an liegen. Die Permutation kann z. B. mit der in Fig. 5 gezeigten Schaltungsanordnung durchgeführt werden. Die Anordnung der Speicher und Multiplexer ist hier jedoch abhängig von dem Transformationsalgorithmus und muß der jeweiligen Aufgabe entsprechend angepaßt werden.

Bezugszeichenliste

D Verzögerungsglied
M Multiplexer
B1 erste Butterfly-Stufe
B2 zweite Butterfly-Stufe
B3 dritte Butterfly-Stufe
PS Permutationsschaltung
P1 - P7 Wichtungsfaktoren
K Korrekturnetz
Sk Skalierung

Patentansprüche

1. Verfahren zur Durchführung von mehrstufigen Transformationen mittels Butterfly-Operationen, **dadurch gekennzeichnet**, daß Stufen im Zeitmultiplex und die anderen Stufen ohne Zeitmultiplex berechnet werden.
2. Verfahren nach Anspruch 1 zur Durchführung von Berechnungen mehrerer Stufen im Zeitmultiplex, **dadurch gekennzeichnet**, daß Ergebnisse nach der Berechnung einer vorhergehenden Stufe permutiert, über eine Verzögerungsleitung in ein Rechenwerk zur Durchführung von Butterfly-Opera-

tionen zurückgekoppelt und einer weiteren Butterfly-Operation unterzogen werden, wobei das Rechenwerk für die Teilberechnungen der Stufen gemeinsam genutzt wird und die Teilberechnungen der Stufen alternierend ausgeführt werden.

3. Schaltungsanordnung zur seriellen Durchführung von mehrstufigen Transformationen im einfachen Zeitmultiplex, **dadurch gekennzeichnet**, daß ein Rechenwerk zur Berechnung einer Butterfly ausgeprägt ist, die Ausgänge des Rechenwerks über eine Schaltung zur Permutation, Verzögerungselemente und Multiplexer an den Eingang des Rechenwerks rückgekoppelt sind, wobei der Signaleingang an den einen Multiplexer und über ein Verzögerungsglied an den anderen Multiplexer geschaltet ist.

4. Schaltungsanordnung nach Anspruch 3 zur Erzeugung eines seriellen Ausgangsdatenstroms, **dadurch gekennzeichnet**, daß die Ausgänge des Rechenwerks zum einen über ein weiteres Verzögerungselement, zum anderen direkt mit einem dritten Multiplexer verbunden sind.

5. Schaltungsanordnung nach Anspruch 4 zur Berechnung einer DCT, **dadurch gekennzeichnet**, daß der Ausgang des dritten Multiplexers mit einem Rechenwerk zur Durchführung einer Butterfly-Operation ohne Zeitmultiplex verbunden ist.

6. Schaltungsanordnung zur Berechnung einer IDCT, **dadurch gekennzeichnet**, daß der Schaltungsanordnung nach Anspruch 3 ein Rechenwerk zur Durchführung einer Butterfly-Operation ohne Zeitmultiplex vorgeschaltet ist.

Hierzu 2 Seite(n) Zeichnungen

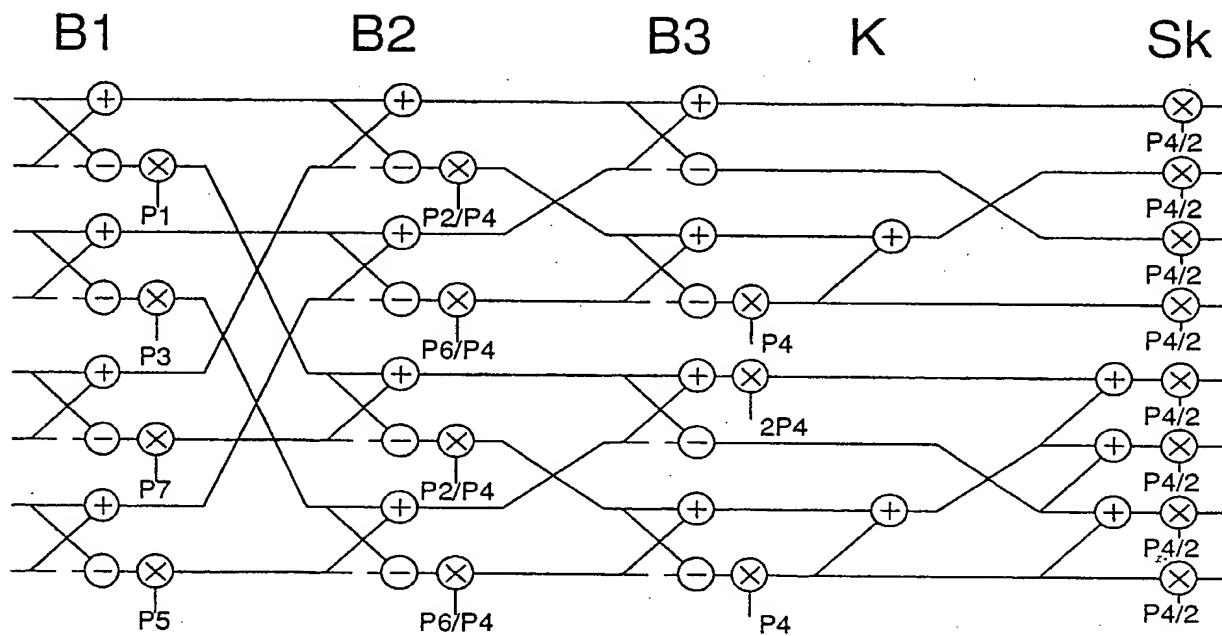


Fig. 1

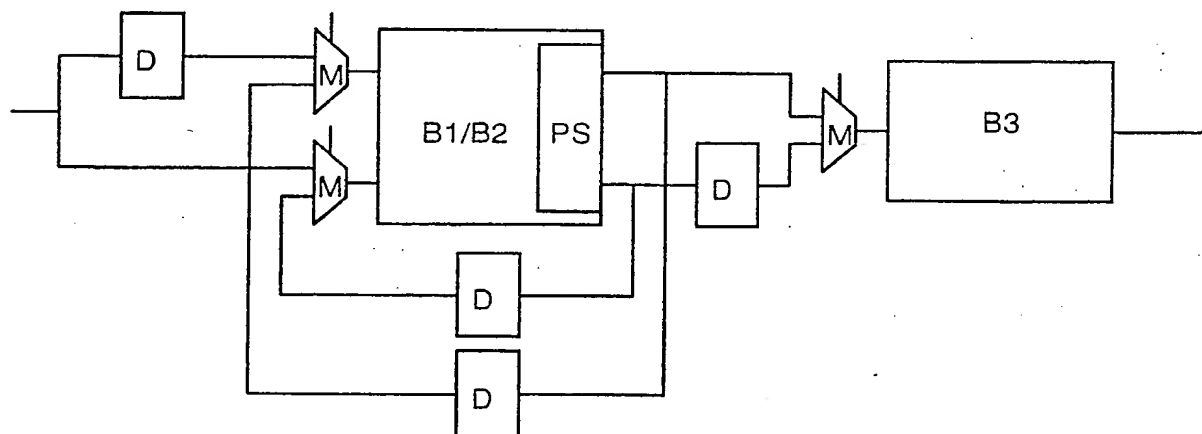


Fig. 2

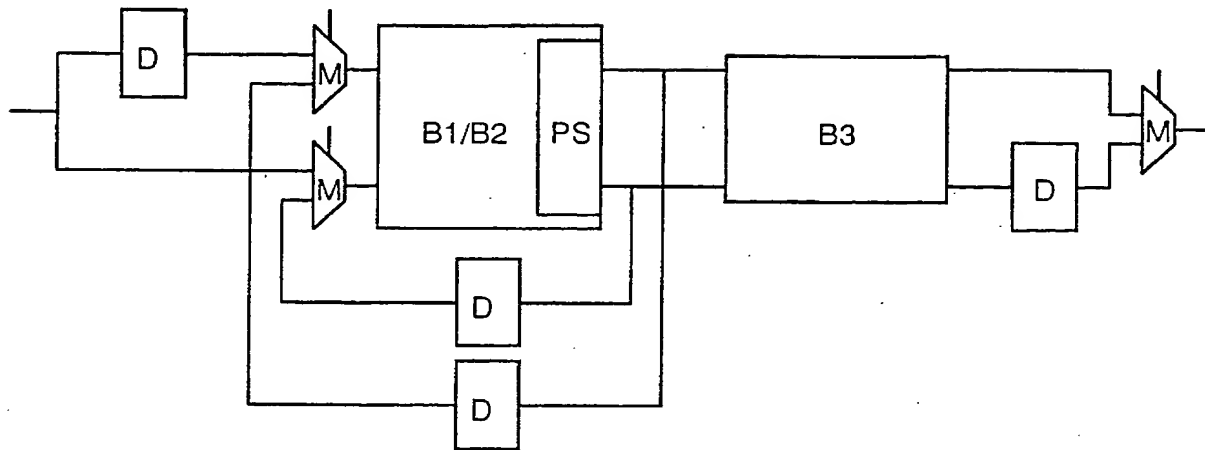


Fig. 3

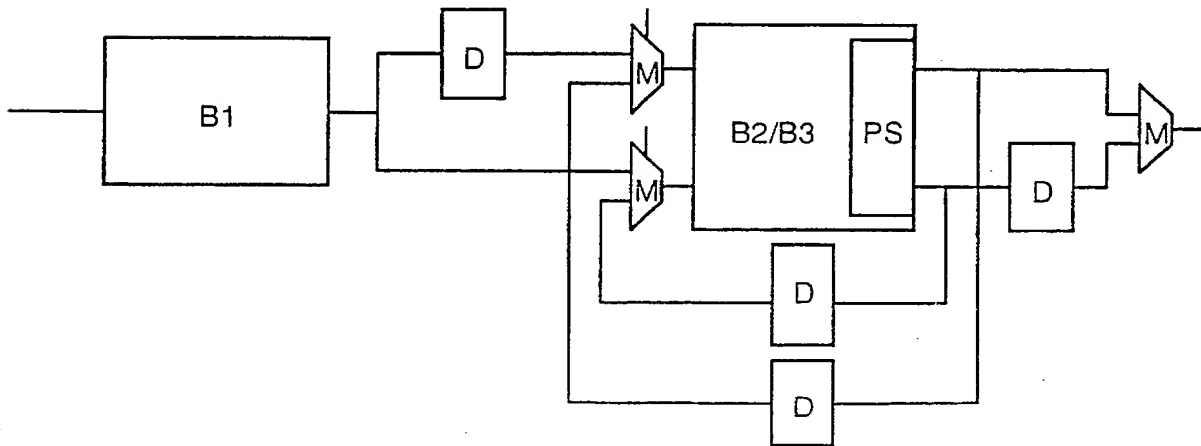


Fig. 4

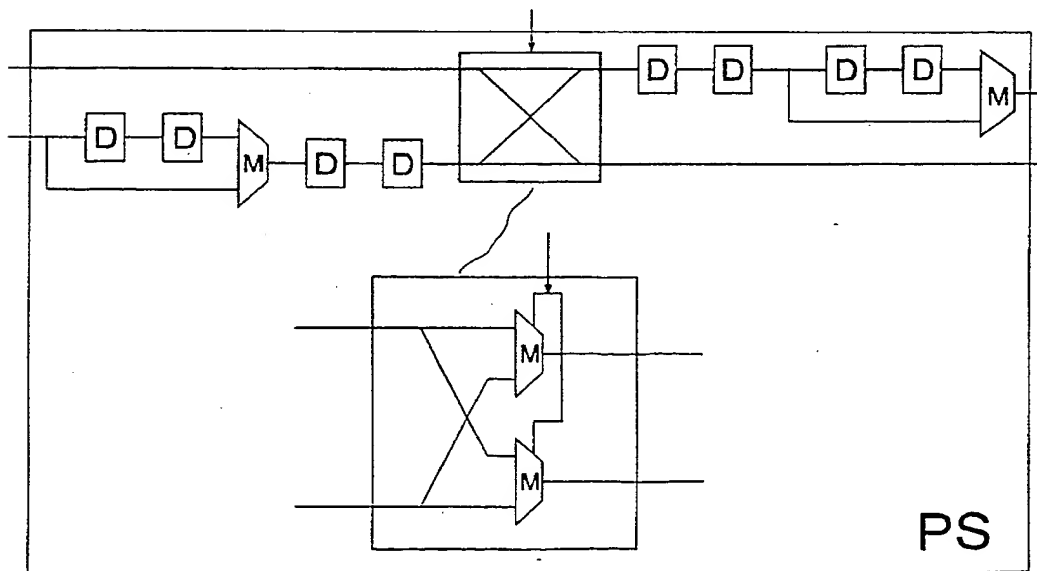


Fig. 5